AD-A107 894 ARMY INVENTORY RESEARCH OFFICE PHILADELPHIA PA F/G 5/1

EVALUATION OF FORECAST METHODS IN AN INVENTORY MANAGEMENT SYSTE--ETC(U)

AUG 81 E P GOTWALS

UNCLASSIFIED IRO-TR-81-5 NL

1 OF 1

AD A107894

END

DATE FILMED

1 82

DTIC

AD-
TECHNICAL REPORT
TR 81-5

EVALUATION OF FORECAST METHODS IN AN INVENTORY MANAGEMENT SYSTEM - AN EMPIRICAL STUDY

U.S. ARMY

INVENTORY
RESEARCH
OFFICE

AUGUST 1981

ROOM 800
U.S. CUSTOM HOUSE
2nd and Chestnut Streets
Philadelphia Pa. 19106

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A107894	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) EVALUATION OF FORECAST METHODS IN AN INVENTORY MANAGEMENT SYSTEM - AN EMPIRICAL STUDY		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) EDWIN P. GOTWALS, III		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS US Army Inventory Research Office Room 800 - US Custom House 2nd & Chestnut Sts., Philadelphia, PA 19106		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS US Army Materiel Development & Readiness Command 5001 Eisenhower Avenue Alexandria, VA 22333		12. REPORT DATE August 1981
		13. NUMBER OF PAGES 47
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Information and data contained in this document are based on input available at the time of preparation. Because the results may be subject to change, this document should not be construed to represent the official position of the US Army Materiel Development & Readiness Command unless so stated.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Demand Forecasts
Evaluation
Forecast Errors
Inventory Simulation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Empirical evidence based on Army wholesale demand data is presented indicating that traditional error measures such as mean squared error and mean absolute deviation do not necessarily relate to performance in an inventory management system. An alternative method is given which performed well when compared with simulation results.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TABLE OF CONTENTS

1. INTRODUCTION AND OVERVIEW

Recent studies (ref [3], [4], [5], [7]) by the Inventory Research Office (IRO) have indicated certain anomalies between methods of evaluating demand forecast algorithms for use in an inventory management system. These studies were based on empirical evidence gleaned from a data base of quarterly demands for over 20,000 aviation repair parts spanning a history of eleven years. In each study both statistical and simulated performance measures were computed. (The statistics were collected by comparing the forecasts with the actual values of the series whereas the simulation results revealed how well the forecasts performed in the simulated supply system. Both methods utilized actual demand data.)

This note emphasizes the problem of inconsistency by showing specific examples where alternative measures of performance yielded different results. In each case a "best" forecast algorithm was sought but discrepancies between the error measures confounded the analysis. Comparative ranks were made using each measure in an attempt to determine the best algorithm(s). It is the rank variability between the evaluation methods that is the concern of this report.

It was hoped that statistical error measures could be found that would be consistent with the supply performance measures. By so doing, both theoretical and empirical guidance would be given for the development of optimal forecasts. The examples cited herein trace the developmental stages of the measures. The first considered were simple statistical expressions such as mean squared error (MSE), mean absolute error (MAD) and average error (BIAS). Based on experiences with these basic measures, percent and relative measures were developed along with changes in processing such as forecast horizon and forecast timing. Next, inventory measures were developed where both cost and performance were measured. The IRO simulator was used to correlate the above measures with inventory performance. Finally a surrogate simulation method was derived to best measure inventory performance.

2. ORGANIZATION

The first seven sections of this report give the basic background details needed to understand the comparative analysis presented in Section 8. The data base common to all the examples is described in Section 3 along with the breakout of the various subsets or stratification classes which were used. In Section 4 the general forms of the forecast algorithms are listed. Sections 5, 6, and 7 describe the three general evaluative methods (statistical error measures, inventory error measures, simulation analysis) and the computational detail of each measure used. The comparative analysis is presented in Section 8. Based on insights gained from Section 8, a surrogate simulation method is developed and tested in Section 9. In Section 10 specific problems in evaluating series with many zeros are mentioned. The conclusions based on the findings are stated in Section 11.

Accession For
NTIS
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A

3. DATA STRUCTURE

The data common to all the examples in this report is from the IRO demand history file which includes 11 years of requisitions and demand by quarter for aviation repair parts. The file contains a sample of 20,865 items from all those in the system in 1966 and subsequently entered. A more detailed description as to the type of item is found in [2].

For the initial analysis each item was classified as low dollar value (LDV) or high dollar value (HDV) according to whether the demand rate averaged over the 11 years was less than or greater than or equal to $50,000; and the requisition rate was less than or greater than or equal to 100 per year. Items with over a million dollars of demand per year were dropped.

The items were further divided into dynamic (DYN) and non-dynamic (NON) based on the Federal Stock Class (FSC). The dynamic components were considered to be those that experience high rotation rates; i.e. rotor blades, transmissions, and turbine engines. For more detail see Cohen [2].

The data breaks out into the following four groups:

HDVDYN	86
HDVNON	262
LDVDYN	1169
LDVNON	19348
	20865

Later on, samples were taken from the data and stratified in classification cells according to the item's requisition frequency and dollar demand. The "strat" class were identified as follows:

Yearly Dollar Demand

Yearly Requisitions	0 - $5000	$5000 - $50000	> $50000
0 - 3	Strat Class 1 N = 335	2 N = 100	3 N = 4
3 - 12	Analyzed 4 N = 124	Analyzed 5 N = 230	6 N = 17
> 12	Analyzed 7 N = 98	Analyzed 8 N = 64	Analyzed 9 N = 115

4. FORECAST METHODS

The five basic forecast algorithms tested were of the form of a weighted moving average (1794), moving median (MED), Kalman filter (KAL), a variable base moving average utilizing a tracking signal (TRIG) and a simple moving average on demand (MOVD). The underlying structure of the algorithms are as follows:

(1794):

This is the current Army forecast procedure which is a weighted eight quarter moving average with weights proportional to quarterly flying hours.

(MEDIAN)

This is the median of the last four quarters.

(Kalman Filter)

For a general class of statistical processes, this is a minimum mean squared error algorithm which is structurally similar to exponential smoothing where the smoothing weights are variable and updated.

(TRIG Tracking Signal)

This is a version of the TRIG methodology which switches between a weighted four and eight quarter moving average where the weights are parameters determining the amount of additional weight to give the current observation.

(MOVD)

This is an eight quarter moving average on demand.

See reference [4] for additional details.

5. STATISTICAL ERROR MEASURES

In an inventory management setting, an optimal forecast is one that performs well over many items each of which may have different forecast horizons (procurement lead time). Each forecast error impacts the overall performance of the system by a different magnitude depending on the demand activity of the series. Various methods of aggregating these errors are presented in this section.

For the experiment, one period (qtr), four period and variable horizons were used for comparative purposes, but for simplicity the notation assumes a one period forecast horizon. There is little in the literature describing empirical work comparing errors between various forecast horizons. Peter Bloomfield of Princeton is doing some theoretical work in this area.

Notation

Let x_{ij} be the demand in the j^{th} quarter (for the i^{th} item) and $\hat{x}_{ij}$ the forecast

F_i = index set of forecasts for i^{th} item i.e. the set of j values for which a forecast is made

η_i = cardinal size of F_i (number of times a forecast was made)

E_{ij} = error $(\hat{x}_{ij} - x_{ij})$, $j \in F_i$

AVG_i = average demand for item i

(Double 12 month moving average starting after the first non-zero demand)

N = number of items

Simple Averages

The first error measures considered were simple averages of traditional measures.

$$MAD = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\eta_i} \sum_{j \in F_i} | E_{ij} |$$

(Mean Absolute Demand)

$$MSE = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\eta_i} \sum_{j \in F_i} (E_{ij})^2$$

(Mean Square Error)

$$BIAS = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\eta_i} \sum_{j \in F_i} E_{ij}$$

(Average Error)

Percent Error Measures

The simple averages give more weight to items with higher demand quantity. Since the items were stratified into homogeneous classes it was desirable to give equal weights to each item in the class, hence the following percent error measures were considered.

$$\text{AVG \% of Forecast} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\eta_i} \sum_{j \in F_i} \frac{|E_{ij}|}{\hat{x}_{ij}} \times 100$$

$$\text{AVG \% of Actual} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\eta_i} \sum_{j \in F_i} \frac{|E_{ij}|}{x_{ij}} \times 100$$

$$\text{AVG \% of Both} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\eta_i} \sum_{j \in F_i} \frac{|E_{ij}|}{\frac{1}{2}(\hat{x}_{ij} + x_{ij})} \times 100$$

Relative Error Measures

Now since many of the series were quite variable, the denominator of the percent error measures did not reflect the steady state demand of the item hence the following relative measures were considered.

$$\text{RELATIVE BIAS} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\eta_i} \sum_{j \in F_i} \frac{E_{ij}}{\text{AVG}_i}$$

$$\text{RELATIVE MAD} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\eta_i} \sum_{j \in F_i} \frac{|E_{ij}|}{\text{AVG}_i}$$

$$\text{RELATIVE MSE} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\eta_i} \sum_{j \in F_i} \left(\frac{E_{ij}}{\text{AVG}_i}\right)^2$$

6. INVENTORY MEASURES

An overforecast error in predicting demand impacts the inventory control system differently than does an underforecast. Overforecasts result in carrying too much stock and increases the possibility of being stuck with obsolete items, whereas underforecasts increase the possibility of not satisfying a customer's orders and in the case of the Army may reduce the readiness of a weapon system. Since there is not a natural tradeoff between these two types of errors, the following separate measures were developed.

Notation

For the given i^{th} demand series $\{x_{ij}\}$ and its corresponding $\{r_{ij}\}$ requisition series (the number of requisitions at quarter j)

let $D_{i,t}(\ell)=\sum_{j=1}^{\ell} x_{i,t+j}$ be the demands over ℓ periods from time t and $\hat{D}_{i,t}(\ell)$, the forecast

$R_{i,t}(\ell)=\sum_{j=1}^{\ell} r_{i,t+j}$ be the number of requisitions over ℓ periods from time t

$$EL_{i,t}(\ell) = (\hat{D}_{i,t}(\ell) - D_{i,t}(\ell)) = \sum_{j=1}^{\ell} (\hat{x}_{i,t+j} - x_{i,t+j})$$ be the errors over lead time ℓ

UP_i = unit price of the i^{th} item

Overforecast Measure

Let
$$OF(i) = \sum_{j\epsilon F_i} \max\left[\frac{EL_{i,j}(8)}{D_{i,j}(8)},0\right] \cdot D_{i,j}(8) \cdot UP_i$$

$$= \sum_{j\epsilon F_i} \max\left\{EL_{i,j}(8),0\right] \cdot UP_i$$

estimate the cost of the extra stock purchased for periods of eight quarters for the i^{th} item.

Then

$$OF = \frac{\sum_{i=1}^{N} OF(i)}{\sum_{i=1}^{N} UP_i \sum_{j \in F_i} D_{ij}(8)} \times 100$$

is an estimate of the percent of the total dollar demand spent on extra stock.*
The base period of 8 is used to indicate the long term effect of an overforecast.

Underforecast Measure

Similarily let

$$UF(i) = \sum_{j \in F_i} \max \left[-\frac{EL_{ij}(PLT)}{D_{ij}(PLT)}, 0 \right] \cdot R_{ij}(PLT)$$

be an estimate of the number of requisitions not satisfied in the procurement lead time, PLT, for the i^{th} item, i.e. if demand is underforecasted, say by 20%, then it is implied that 20% of the requisitions will not be satisfied. The reason requisitions are used instead of demand quantity is because of DoD policy [DoDI 4140.39].

Then

$$UF = \frac{\sum_{i=1}^{N} UF(i)}{\sum_{i=1}^{N} \sum_{j \in F_i} (R_{ij}(PLT))} \times 100$$

is an estimate of the percent of the total requisition not satisfied over all the items.

The procurement lead time of the item is the quickest time stock could be replenished after a new order is placed. Hence, in an underforecast situation, the reorder point will probably be crossed within a PLT, and a new forecast would be made.

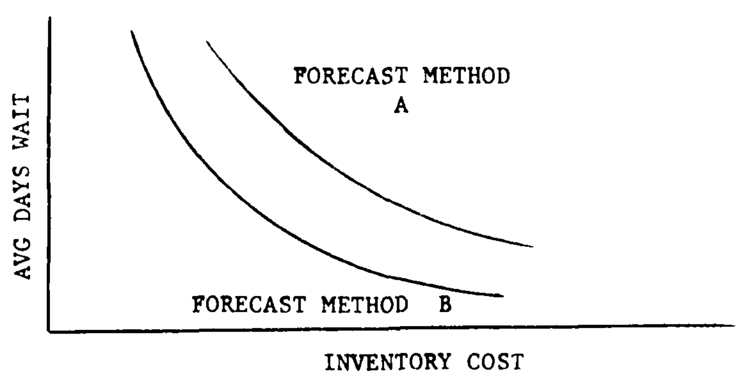
* This is extra stock for the forecast period and the cost estimate associated with it does not consider its utility as safety stock for future underforecasts.

7. SIMULATION ANALYSIS

The third way to evaluate the forecast algorithms is to observe their performance in the Army's wholesale management system. This is done by using the DoDI 4140.39 Simulator which was designed to compare the performance of alternative policies within this management system. (Details of this simulation may be found in [1], [2], and [4].) The basic operational sequence of the simulator can best be described from the following excerpt from [1].

The demand history for a given item is fed both to the forecast algorithm process, which computes the demand forecast for each period; and to the asset monitor, which reduces on-hand stock - thus possibly triggering an order for replenishment - and computes the final performance estimates. The demand forecasts are fed to the inventory model, which computes the reorder point needed for asset monitoring and the order quantity needed when orders are placed. The final performance estimates are averaged over the simulated time for each item, and then aggregated across all items. This is done separately for each forecasting algorithm.

The method used to compare policies is to plot a cost-performance curve for each rule being considered. These curves are traced through several "λ" points where the lambda (λ) value reflects the cost of a backorder. (The lambda value is set prior to a simulation run). An increase in λ results in an increase in inventory cost and an improvement in performance, hence λ can be considered as a control knob in determining various points on the curve. A typical graph comparing two policies is shown below:



(Note: The "Inventory Cost" as shown on the graph includes the cost to order and the cost to hold averaged over all the items and the number of simulated years. The performance measure, "Average Days Wait," is a weighted average of the average wait per requisition per item.)

8. COMPARISON OF ERROR MEASURES

It was hoped that a simple statistical error measure could be found that ranked the forecast algorithms according to their performance in an inventory management system. By so doing, both theoretical and empirical guidance would be given in the development of optimal forecasts. Currently such measures as mean squared error and mean absolute deviation are minimized but these criteria do not necessarily guarantee optimal performance for inventory management as shown in this section.

The following is a list of observations made from a sequence of experiments comparing error measures while using the forecast algorithms introduced in Section 4. Tables found in the Appendix are referenced after each observation.

Initial Observations:

a. The simple averages MAD, MSE, and BIAS gave more weight to higher demand items. This was confirmed by reviewing individual item contribution to the aggregation. Refer to Appendix A1.

b. The average percent error for both forecast and actual yielded inconsistencies between their 1 quarter and 4 quarter measures and also between themselves. This was felt to be due to the variability of their denominators. Refer to Appendix A2.

c. Inconsistency still plagued the rankings of the relative error measures. Refer to Appendix A3.

d. Simulation results did not correspond with statistical results where a consensus was used to rank the statistical measures. Refer to Appendix A4.

Procedural Changes

Based on the problems between the statistical measures plus the desire to predict performance in an inventory setting, the following procedural changes were made to the method of computing the statistical error measures.

a. Forecast only after a non-zero demand; this is the only time a reorder point may be triggered and where the forecast is actually used (alternative would be periodic review which wasn't considered).

b. Use only the item's PLT as a forecast horizon; again this more clearly corresponds to the inventory model.

c. Use a constant, 8 quarter PLT, for the basis of the overforecast error measure.

Additional Observations:

Applying the above process changes to both the statistical and inventory error measures and to the nine stratification classes (Section 3) and comparing these results with the simulator yielded the tables in Appendix A5. (Note the large positive bias in many cases.) The following inferences were drawn:

a. Limiting the algorithms to those methods not exhibiting a large positive bias, several comments can be made.

(1) The simulator favored those methods with small positive bias. Additional experiments confirmed this. (Forecasts were increased on methods that were unbiased and the simulated performance improved; refer to Appendix A6).

(2) The inventory measure for underforecast compared well with the simulator but this measure can be made to converge to zero by making overforecasts; hence the need to limit the positive bias.

b. Consistency improved between the ranks of the various statistical measures, particularly REL MSE and REL MAD.

Simulation Experimentation:

Experimentation with the simulator revealed the following:

a. Removing the phased deliveries and the R and Q constraints* eliminated the simulator's preference for positive bias forecasts. (Refer to A7)

b. The simulator performance was sensitive to individual items with high demand whereas the error measures were designed to give equal weight to each item. For an example refer to Appendix A8.

Final Adjustments:

In a final effort to correlate the statistical measures with the simulator the following adjustments were made:

a. Screen from the simulation results those items that carried extreme high cost or poor performance of specific algorithms. (These items would be analyzed separately).

---

*In actual practice the R and Q decision parameters are constrained to make them less sensitive to unreasonable fluctuations of the data. Phased deliveries are used to spread delivery over several months after the PLT. Removing both of these conditions made the simulator more consistent to the model used to optimize the parameters.

b. For use with the simulator and error measures, use a simple 8 quarter average for the forecast if the item had been inactive for year prior to the demand triggering the forecast; this would handle the migration of an item from an active strat class to an inactive one without unduly penalizing the algorithm which would normally work well in an active class and does poorly in the less active class.

Final Observations:

The results of these changes for the five active stratification classes are in Appendix A9. Looking at the rank statistics the following inferences can be made:

a. For the most part, the ranks between the 1 quarter and PLT forecast horizons agreed.

b. The ranks between the three statistical measures (MAD, REL MAD, REL MSE) were more consistent than before.

c. There were still some inconsistencies between the statistical measures and the simulator.

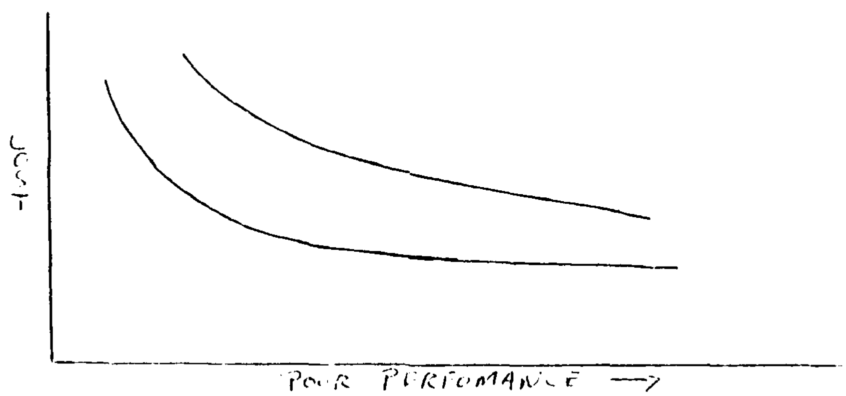
9. SURROGATE SIMULATOR

Since the underforecast measure compared well with the simulator, when the overforecasts were constrained, the following method was used to measure the combined effects of both over and underforecasts.

The underforecast measure is used as a performance indicator. It is an estimate of the percent of the total number of requisitions not satisfied while using a specific forecast method.

Similarly the overforecast measure is used as a cost indicator. This is an estimate of the percent of the total dollar demand spent on extra stock not used to satisfy demands.

By adjusting the forecast (adding or subtracting a percent of the forecast) to represent the impact of adding safety stock (negative or positive), various performance levels are attained. With each level of performance a corresponding cost is computed yielding a cost vs performance curve for each forecast method.



Curves closest to the X-Y axis represent better performance or lower cost.

To test this concept, the three strat classes ST9, ST8, and ST7 were used. Analysis of the simulation results revealed that of the five forecast algorithms tested only MOVD performed differently from the others. This occurred in classes ST9 and ST8 where MOVD performed worse than the others.

Using the surrogate measure with the three strat classes, again MOVD appeared to be the worse performing algorithm in ST9 and ST7 strat classes. Refer to Appendix A10 for results. To date this surrogate simulation method has had the most consistent performance when compared with the simulator.

10. INACTIVE ITEMS

A major problem in statistically comparing different forecast algorithms using series with many zero observations is that traditional error measures give substantial weight to deviations from these zero modal values. Consequently, these measures favor algorithms yielding routinely small forecasts and, in fact, will probably most favor the degenerate algorithm which forecasts zero all the time.

Appendix A3 of [3] describes results from an IRO study that supports the above remarks. In particular, it demonstrates that MAD did favor the degenerate forecasts of zero. Additional work done on various other studies at IRO also support these findings. Appendix A11 of this report gives additional credence to this claim using samples of low demand items from the subject data base.

There are also problems with using the simulator for inactive items. The simulator doesn't adequately duplicate the Army's policy of managing low demand items; details of which will not be discussed here. A problem which is more cumbersome to resolve deals with the way the simulator distributes the quarterly data into 32 periods per year. This "data enrichment" procedure was not designed to handle sparse demands and severely distorts the actual series. [9]

The only resolution to this problem to date has been to use the statistical method of Section 9. A further study by IRO will deal with the management of low demand items and at that time a simulator specifically designed for these items will be developed.

11. CONCLUSION

By showing empirical anomalies between the various error measures this paper has brought to light a question which is often overlooked when addressing a forecast problem, i.e., what is being optimized? There is much in the literature on minimizing the one period ahead mean squared error, but little has been said about errors over a long term horizon and their impact on the working environment.

The measure given in Section 9 is an alternative to the classical error measures which relates errors to performance in an inventory management system. This method is simpler to program and less expensive to run then a detailed simulation and its results have compared well with the IRO inventory simulator.

APPENDIX A1

INDIVIDUAL ITEM MEASURES

DATA: HDVDYN , N = 86

FORECAST METHOD: TRIG

COMMENTS: **Item 10's contribution to the aggregation was much larger than the other items for measures MAD, MSE, and BIAS.**

ITEM	MAD	MSE	MAPE	MAD/I	MAD/.SCALE	BIAS
10	470.086	375420.07	.317	.247	.267	207.[illegible]
10	470.086	375420.07	.317	.247	.267	207.[illegible]
20	1.131	1.56	.747	.514	.564	[illegible]
20	1.131	1.56	.747	.514	.564	[illegible]
30	15.297	311.16	.383	.344	.342	[illegible]
30	15.297	311.16	.383	.344	.342	[illegible]
40	3.974	32.27	.602	.359	.400	[illegible]
40	3.974	32.27	.602	.359	.400	[illegible]
50	1.361	2.98	.722	.919	.626	[illegible]
50	1.361	2.98	.722	.919	.626	[illegible]

APPENDIX A2

INCONSISTENCY WITH % ERROR MEASURES

DATA: HDVDYN , N = 96

FORECAST METHOD: Seven Different

COMMENTS: The range in the ranks are too variable to make a reasonable decision.

		MAD/F		MAD/(A+F)/2		
		1qtr	4qtr	1qtr	4qtr	
MED		2.46	5.19	.875	.902	
TRIG		1.96	2.74	.892	.812	
KAL -	(MA3)	1.83	3.23	.878	.809	
1194		1.89	2.22	.911	.781	
KAL -	(NEW)	1.54	2.57	.885	.762	
KAL -	(MA2)	1.83	3.68	.871	.828	
REGKB	(KAL TYPE)	2.12	3.54	.891	.830	
		← RANKS →				RANGE
MED		7	7	2	7	5
TRIG		5	3	6	4	3
KAL	(MA3)	3	4	3	3	1
1194		4	1	7	+	5
KAL -	(NEW)	1	2	4	1	3
KAL -	(MA2)	2	6	1	5	5
REGKB	(KAL TYPE)	6	5	5	6	1

APPENDIX A3

INCONSISTENCY WITH REL ERROR

DATE: HDVDYN , N = 86

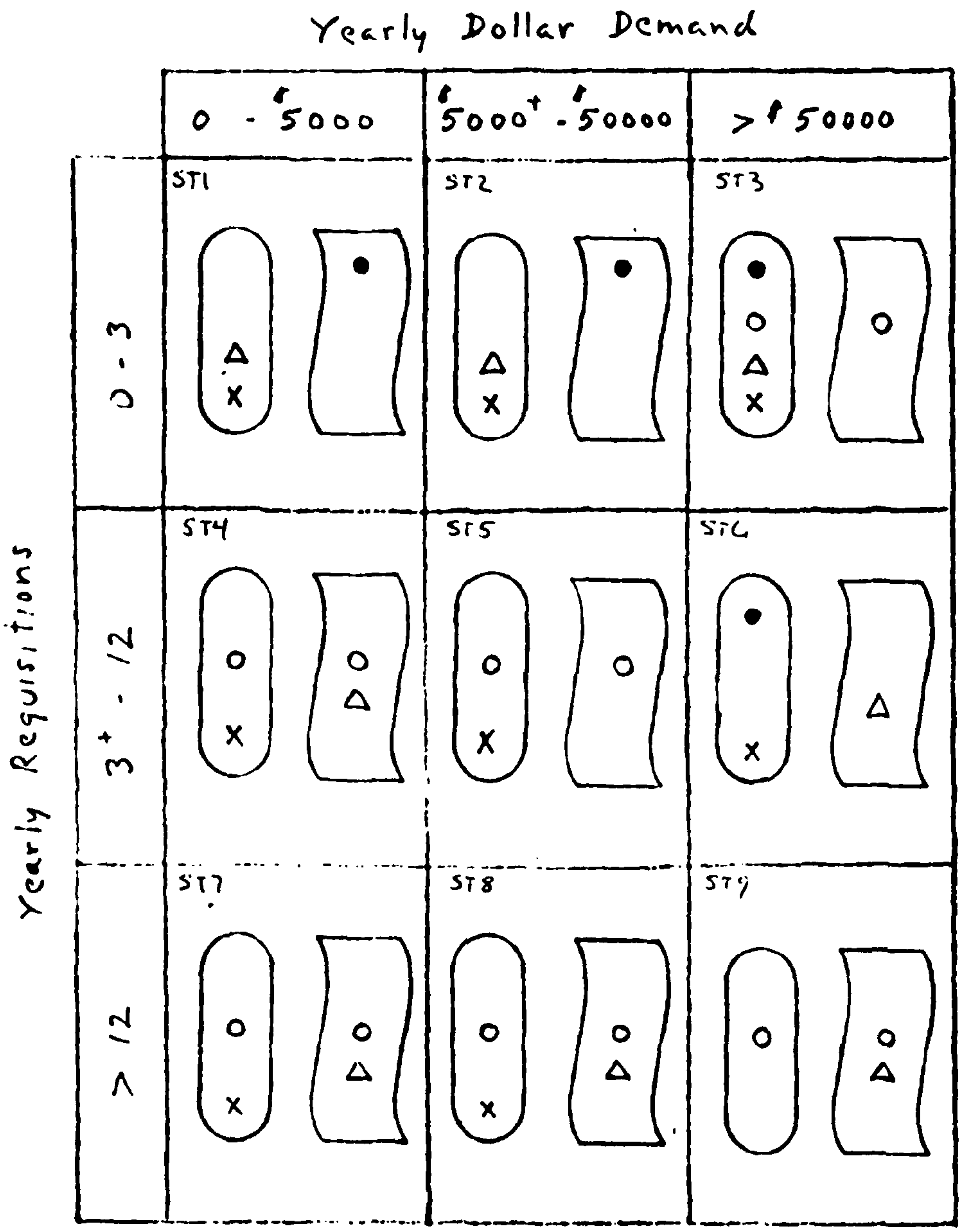
FORECAST METHOD: Seven Different

COMMENTS: Variability between ranks of Relative Error Measures.

		REL	MAD	REL	MSE	
		1 qtr	4 qtr	1 qtr	4 qtr	
MED		.244	.805	.163	3.05	
TRIG		.259	.837	.174	3.19	
KAL -	(MA3)	.254	.802	.149	1.89	
1794		.260	.777	.146	1.19	
KAL -	(NEW)	.255	.779	.147	1.74	
KAL -	(MA2)	.254	.807	.152	1.81	
REGKB	(KAL TYPE)	.257	.808	.153	1.34	
			← RANKS →			RANGE
MED		1	6	4	6	5
TRIG		6	7	7	7	1
KAL	(MA3)	2½	3	3	5	2
1754		6	1	2	1	4
KAL	(NEW)	4	2	1	3	3
KAL	(MA2)	2½	4	5	4	2½
REGKB	(KALTYPE)	5	5	6	2	4

APPENDIX A4

STATISTICAL RESULTS VS SIMULATION RESULTS



Pattern of best algorithms: 1794 ●
by statistics 0 and by KALI o
simulation ▯ TRIG Δ
MED4 x

APPENDIX A5

STATISTICAL AND INVENTORY ERROR MEASURES USING PROCESS CHANGES TO BETTER EMULATE INVENTORY MANAGEMENT

(Actual Values and Comparitive Ranks Along With Simulation Ranks)

DATA: STRAT CLASSES ST4, ST5, ST6, ST7, ST8

FORECAST METHODS: TRIG, MEDIAN, 1794, KALMAN

STATISTICAL AND INVENTORY ERROR MEASURES

		ERROR	REL ERROR	MAD	REL MAD	MSE	UNDER	OVER	# UNDER
	TRIG	1.631	.144	10.009	.845	1.389	.241	.454	224
ST4	MMED	-1.212	-.166	9.629	.809	1.360	.329	.295	275
62	1794	.703	.025	9.946	.826	1.395	.271	.450	247
	KALM	1.040	.079	9.399	.786	1.248	.243	.370	224
	KALM	10.332	.197	22.378	.792	1.195	.257	.460	776
ST5	MMED	-.865	-.048	20.304	.796	1.152	.367	.366	960
230	1794	5.323	.135	20.464	.809	1.277	.278		
	TRIG	6.616	.229	21.586	.858	1.311	.272		
	TRIG	15.618	.238	181.529	1.032	3.305	.242	.657	56
ST6	MMED	-65.672	-.113	112.573	.812	1.844	.290	.354	54
17	1794	37.381	.324	167.688	1.064	2.893	.277		64
	KALM	53.889	.318	130.469	1.022	3.155	.226		55
	KALM	-10.384	-.059	34.605	.564	.524	.297	.256	276
ST7	MMED	-13.457	-.119	35.083	.601	.635	.327	.233	299
49	1794	-10.311	-.069	38.121	.622	.633	.318		281
		-7.593	-.011	34.319	.586	.606	.275		277
	TRIG	3.808	.005	103.965	.349	.214	.153	.182	318
ST8	MMED	-6.227	-.051	104.57	.348	.210	.176	.154	416
64	1794	-22.394	-.032	107.121	.377	.262	.198		705
	KALM	-19.277	-.030	100.84	.218	.181	.182		393

COMPARATIVE RANKS

DATA BASE	FORECAST	ERROR	REL ERROR	MAD	REL MAD	REL MSE	UNDER	SIMULATION	# UNDER
62 ST4	1794	1	1	3	3	3	3	2	247
	MED	3	3	2	2	3	4	3	275
	TRIG	4	4	4	4	3	1.5	1	724
	KALMAN	2	2	1	1	1	1.5	4	224
230 ST5	1794	2	2	1.5	2	3	2.5	3	
	MED	1	1	1.5	2	1	4	2	960
	TRIG	3	4	3	4	4	2.5	4	
	KALMAN	4	3	4	2	2	1	1	774
17 ST6	1794	2	3.5	2	3	2	3		64
	MED	4	1	1	1	1	4		54
	TRIG	1	2	3.5	3	3.5	2		56
	KALMAN	3	3.5	3.5	3	3.5	1		55
49 ST7	1794	2.5	3	4	4	3.5	3	$2\frac{1}{2}$	231
	MED	4	4	3	3	3.5	4	4	299
	TRIG	1	1	1.5	1.5	2	1	$2\frac{1}{2}$	277
	KALMAN	2.5	2	1.5	1.5	1	2	1	274
64 ST8	1794	4	1.5	4	4	4	4	$3\frac{1}{2}$	405
	MED	2	4	2.5	2.5	2.5	2	$3\frac{1}{2}$	416
	TRIG	1	1	2.5	2.5	2.5	1	1	315
	KALMAN	3	2.5	1	1	1	3	2	398

APPENDIX A6

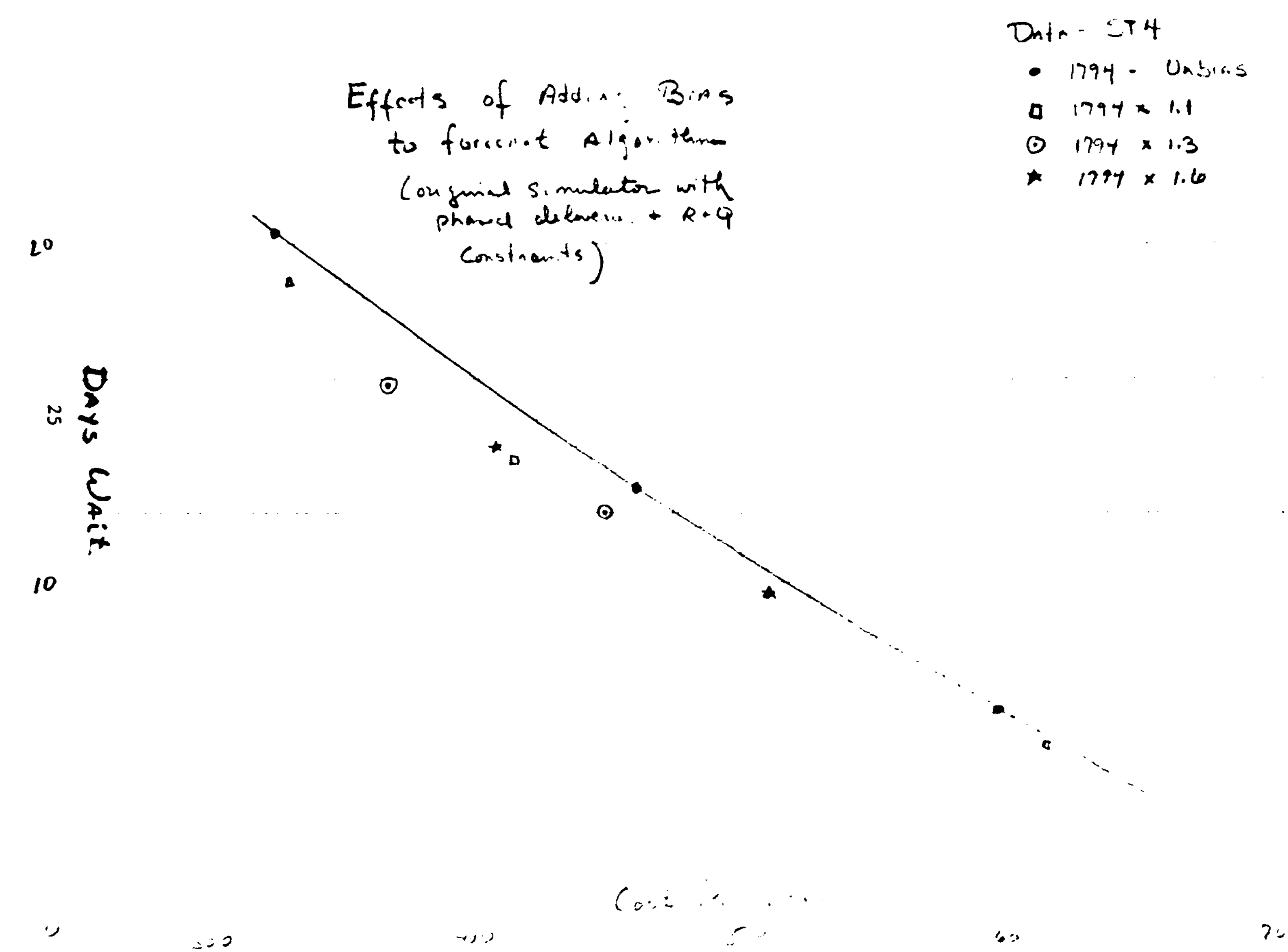
EFFECTS OF ADDING BIAS TO FORECAST METHODS USING SIMULATOR

WITH PHASED DELIVERIES AND R-Q CONSTRAINTS

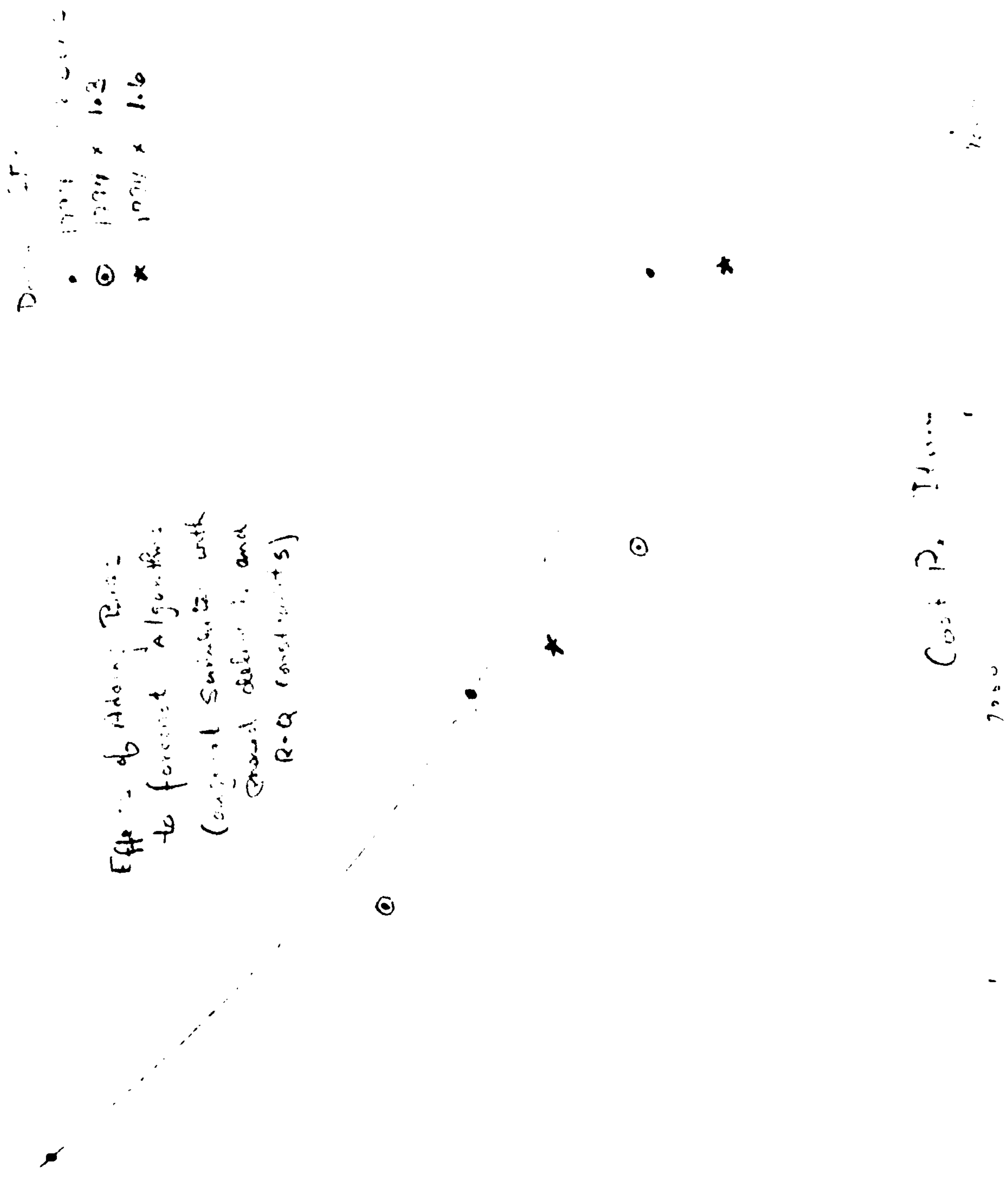
DATA: STRAT CLASSES ST4 AND ST8

FORECAST METHOD: 1794 ADJUSTED FOR BIAS

Appendix A6 (pg 1 of 2)



Appendix A6 (Pg 2 of 2)



DAYS WAIT

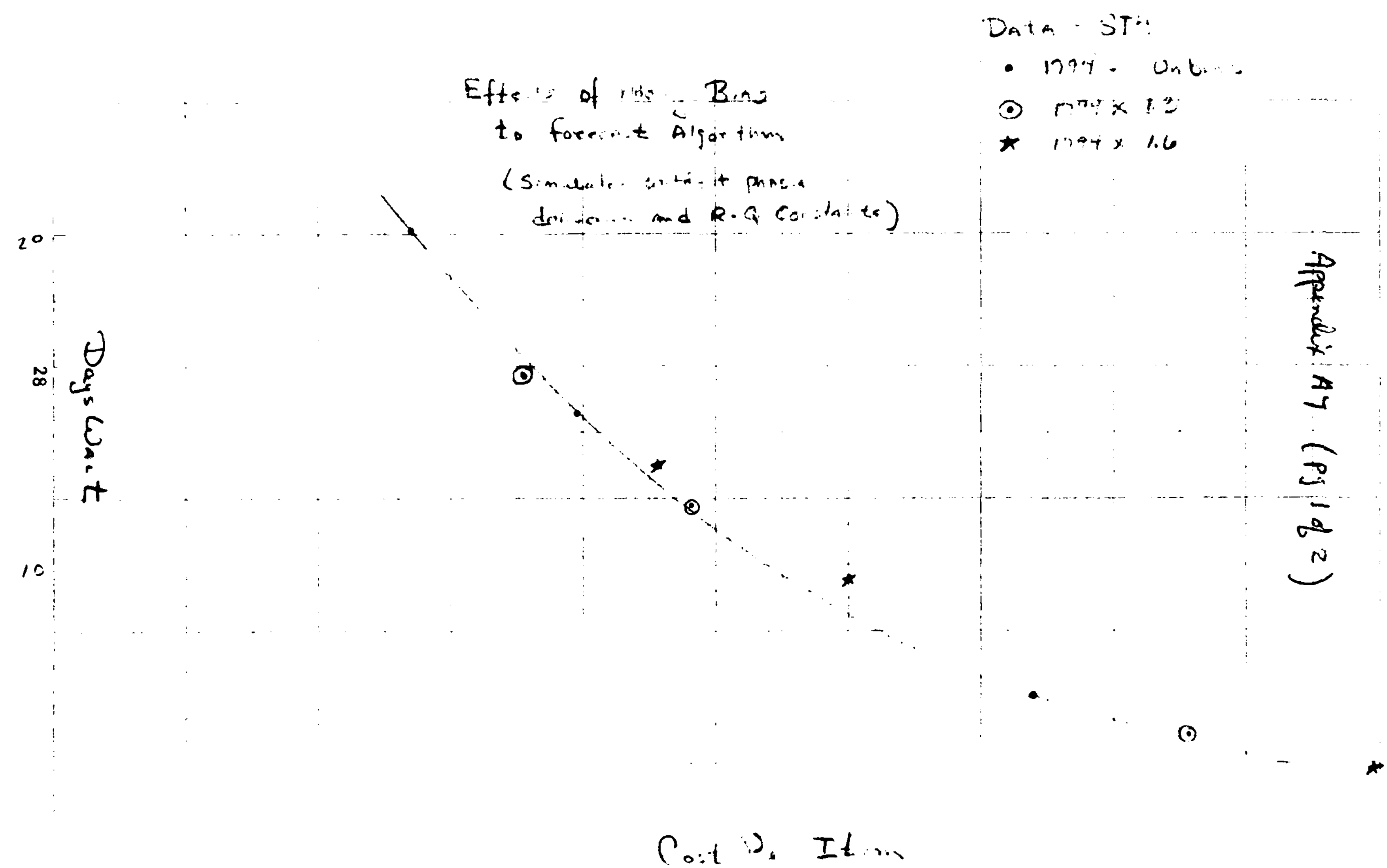
APPENDIX A7

EFFECTS OF ADDING BIAS TO FORECAST METHOD USING SIMULATOR WITHOUT BOTH PHASED DELIVERIES AND R-Q CONSTRAINTS

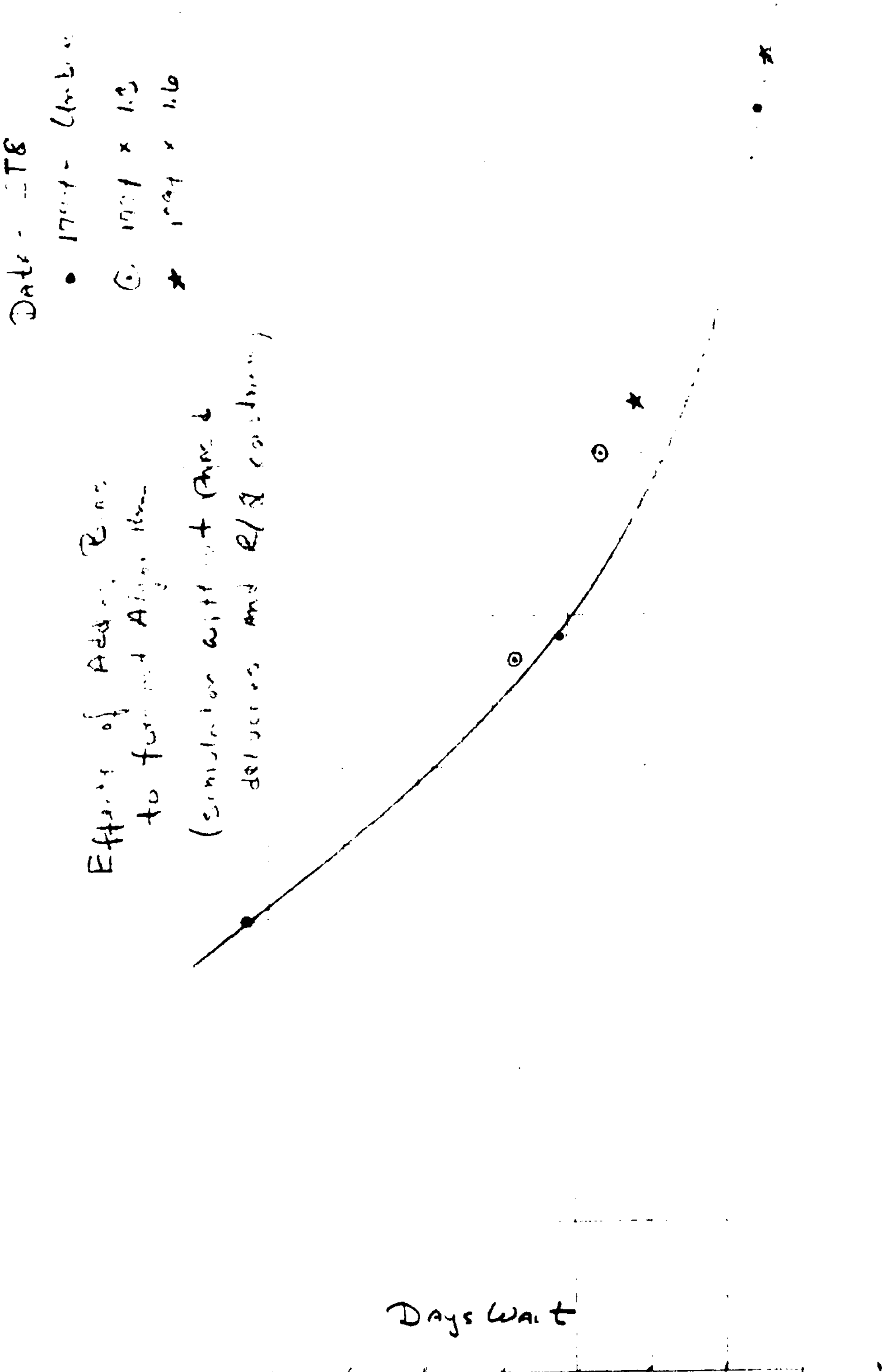
DATA: STRAT CLASS ST4 AND ST8

FORECAST METHOD: 1794 ADJUSTED FOR BIAS

Appendix A7 (pg 1 of 2)



Appendix A7 (Pg 2 of 2)



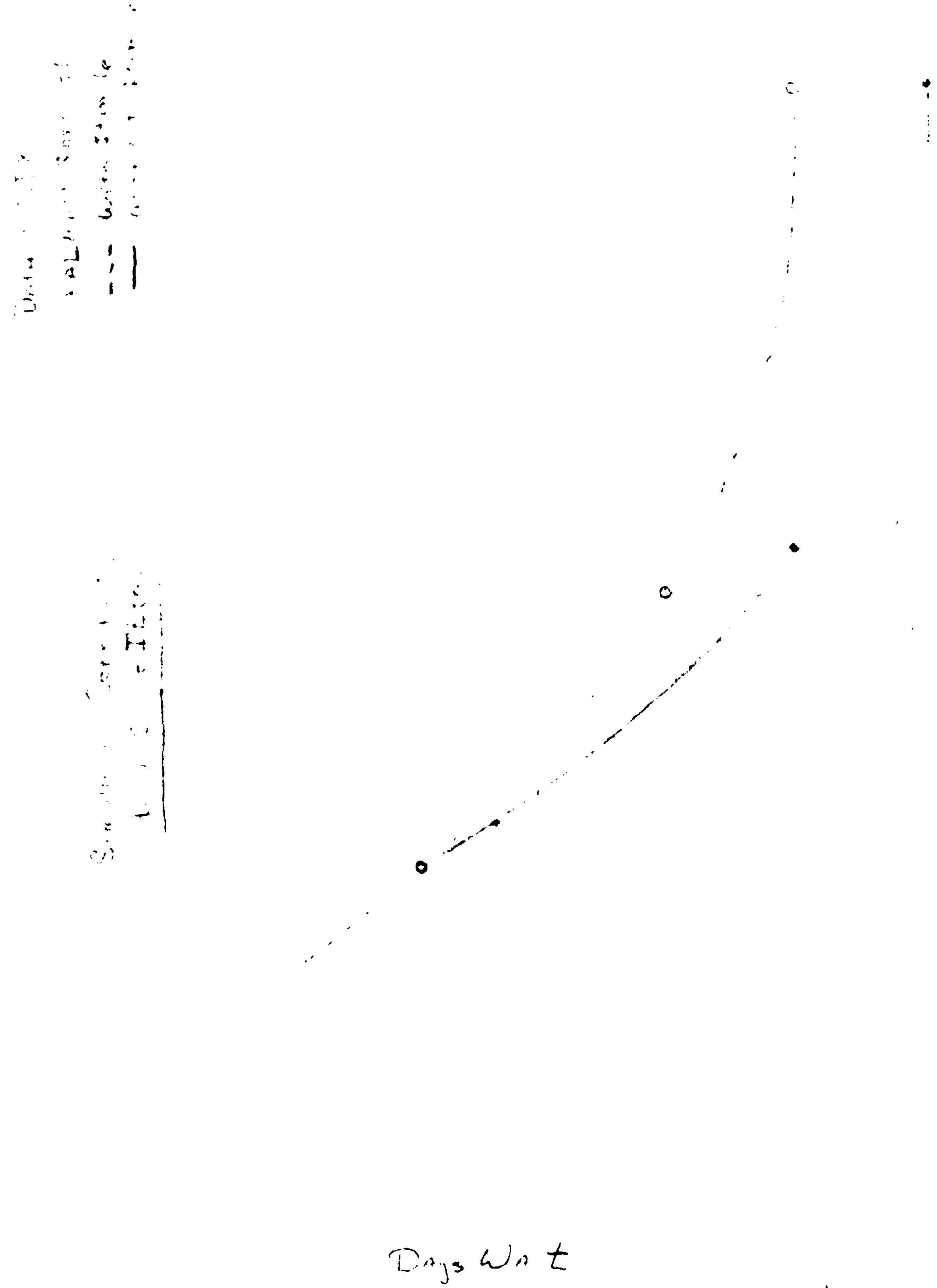
APPENDIX A8

EXHIBIT OF SIMULATOR SENSITIVITY TO A SINGLE ITEM

DATA: STRAT CLASS ST8

FORECAST METHOD: KALMAN FILTER WITH AND WITHOUT ITEM #6

Appendix A8



APPENDIX A9

FINAL COMPARISON OF STATISTICAL ERROR MEASURES AND SIMULATION RESULTS

STRAT CLASSES: ST4, ST5, ST7, ST8, ST9

ERROR MEASURES: ERROR (BIAS), REL ERROR, MAD, REL MAD, REL MSE, SIMULATION RANKINGS

		ERROR		RELATIVE ERROR		MAD		REL MAD		REL MSE	
		1st gtr	PLt	1st gtr	PLt	1st gtr	PLt	1st gtr	PLt	1st gtr	PLt
ST4	1791	-.117	.717	-.047	.028	4.625	9.910	1.032	.823	2.301	1.384
	KAL	.021	1.178		.077	4.506	9.414	1.00	.790	2.086	1.247
	TRIG	.342	1.968	.096	.171	4.722	9.984	1.050	.858	2.707	1.424
	MED	-.555	-.369	-.111	-.031	4.514	9.471	1.024	.806	2.281	1.313
	MOD	.453	2.435	.045	.137	4.98	10.95	1.08	.89	2.54	1.642
ST5	MH	.615	5.323	.053	.137	2.446	20.494	1.060	.810	2.624	1.278
	KAL	1.249	8.765	.093	.155	7.882	22.950	1.056	.792	2.607	1.150
	TRIG	.957	6.54	.187	.242	2.486	21.754	1.162	.865	2.725	1.384
	MED	-1.128	-.019	-.032	.025	7.246	20.022	1.022	.778	2.500	1.107
	MOD	.743	6.204	.103	.200	7.75	21.031	1.084	.843	2.473	1.317
ST7	1791	-2.880	-10.314	-.072	-.069	15.593	38.124	.731	.622	1.025	.630
	KAL	-2.667	-10.335	-.057	-.056	14.483	34.651	.692	.566	.912	.531
	TRIG	-1.778	-7.477	-.011	-.009	14.291	34.223	.690	.583	.926	.601
	MED	-3.838	-13.345	-.114	-.116	14.639	34.971	.703	.599	1.017	.629
	MOD	-.824	-4.090	-.008	.006	15.460	37.243	.735	.616	1.038	.630
ST8	D14	-6.319	-22.399	-.020	-.032	44.61	107.121	.443	.377	.373	.262
	KAL	-4.657	-17.798	-.020	-.031	41.57	100.45	.409	.344	.312	.213
	TRIG	1.594	3.303	.013	.005	42.61	103.96	.409	.349	.310	.214
	MED	-1.626	-6.227	-.037	-.051	42.649	104.571	.410	.348	.309	.210
	MOD	-.201	.033	.049	.057	43.98	110.12	.474	.424	.471	.460
ST9	1791	1.108	-4.619	-.009	-.055	23.49	58.25	.457	.373	.47	.274
	KAL	.358	-5.438	-.019	-.053	22.069	51.25	.473	.371	.440	.261
	TRIG	2.064	2.361	.027	.004	21.57	53.53	.454	.389	.453	.279
	MED	-.778	-6.96	-.025	-.059	22.198	57.02	.490	.407	.482	.310
	MOD	2.521	1.902	.023	-.009	23.478	61.698	.489	.402	.476	.310

COMPARATIVE RANKINGS

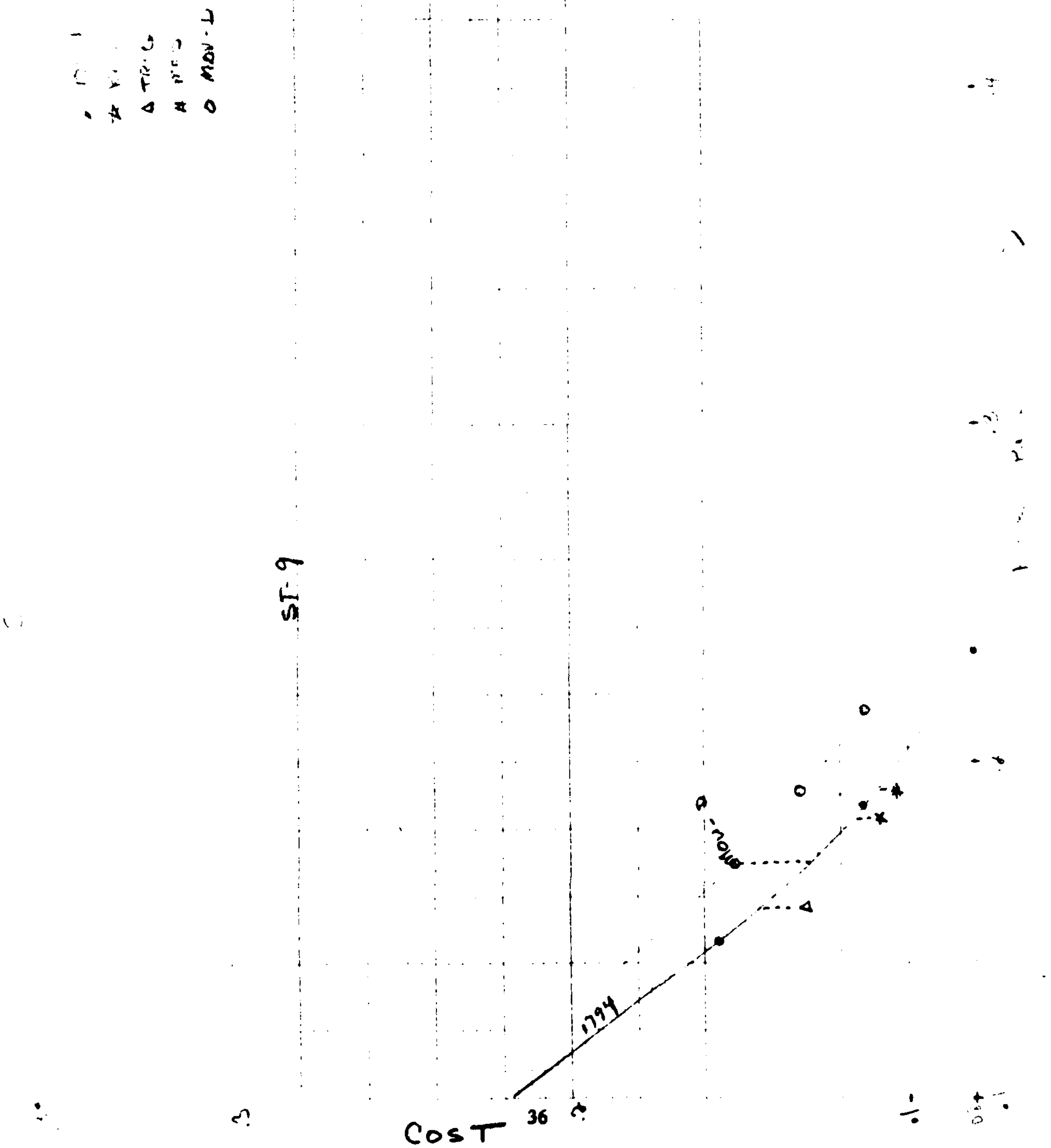
Error Measures Computed	Inventory Types: Over	Under	ERROR 1st	PLT	REL ERROR 1st	PLT	MAD 1st	2nd	REL MAD 1st	PLT	REL ASE 1st	2nd	[illegible]	[illegible]	SIMULATOR ASE	ALL
1794	.450	.271	2	2	1	1	[illegible]		3	3	4	3	3	4		2
KAL	.370	.243	1	3		3	1	1	1	1	1	1	2	2		4
TRIG	.454	.241	3	4	3	5	4	4	4	4	2	4	4	1		1
MED	.295	.329	5	1	4	2	2	2	2	2	3	2	1	5		5
MOD	.591	.245	4	5	2	4	5	5	5	5	5	5	5	3		3
(ST4)																
1794		.278	1	2	2	2	1	2	3	3	3	3		4		3
KAL	.460	.257	5	5	3	3	5	5	2	2	2	2		1		2
TRIG		.272	3	4	5	5	2	3	5	5	5	5		3		4
MED	.366	.367	4		1	1	3	1	1	1	1	1		5		1
MOD	.542	.263	2	3	4	4	4	4	4	4	4	4		2		5
(ST5)																
1794	.280	.318	4	3	4	4	5	5	5	5	4	4	3	4		5
KAL	.256	.296	3	4	3	3	2	2	2	1	1	1	2	3		3
TRIG	.285	.273	2	2	2	2	1	1	1	2	2	2	4	2		1
MED	.233	.327	5	5	5	5	3	3	3	3	3	3	1	5		4
MOD	.354	.268	1	1	1	1	4	4	4	4	5	5	5	1		2
(ST7)																
1794	.202	.198	5	5	3	3	5	4	4	4	4	4	4	5		3
KAL	.172	.182	4	4	2	2	1	1	1	1	3	2	2	4		4
TRIG	.182	.153	2	2	1	1	2	2	2	2	2	3	3	1		1
MED	.154	.176	3	3	4	4	3	3	3	3	1	1	1	3		2
MOD	.319	.172	1	1	5	5	4	5	5	5	5	5	5	2		5
(ST8)																
1794	.114	.183	3	3	1	4	5	4	3	2	3	2	5	4	3	3
KAL	.112	.181	1	5	2	[illegible]	2	2	1	1	1	1	4	3	1	2
TRIG	.130	.157	4	2	4	1	1	1	2	3	2	3	1	[illegible]	2	1
[illegible]	.103	.157	2	4	5	5	3	3	5	5	5	5	3	5	4	
[illegible]	.132	.170	5	1	3	2	4	5	4	4	4	4	2	2		4
[illegible]																

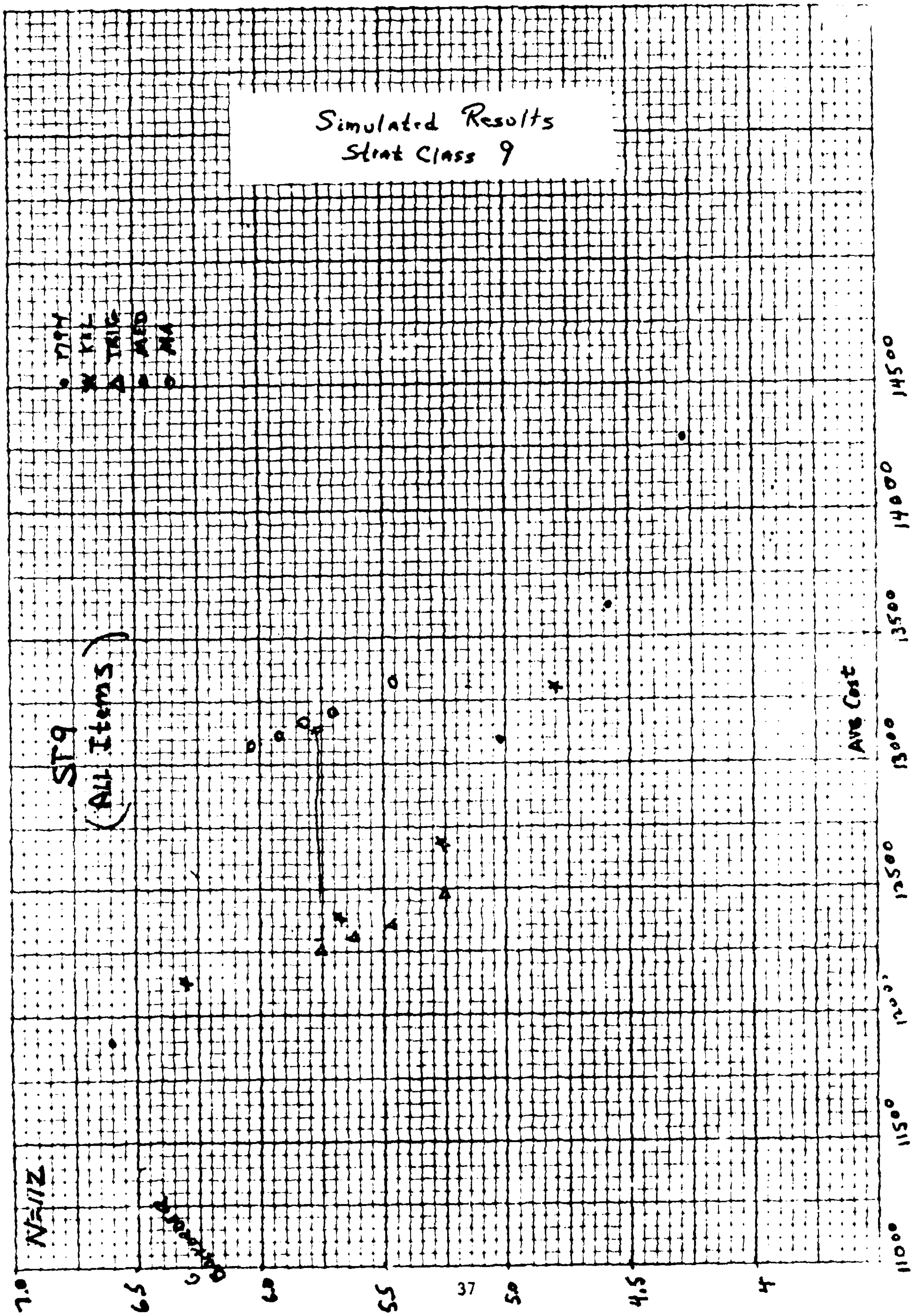
APPENDIX A10

GRAPHS COMPARING COST VS PERFORMANCE RELATIONSHIPS USING STATISTICAL (SURROGATE SIMULATOR) AND SIMULATED MEASURES

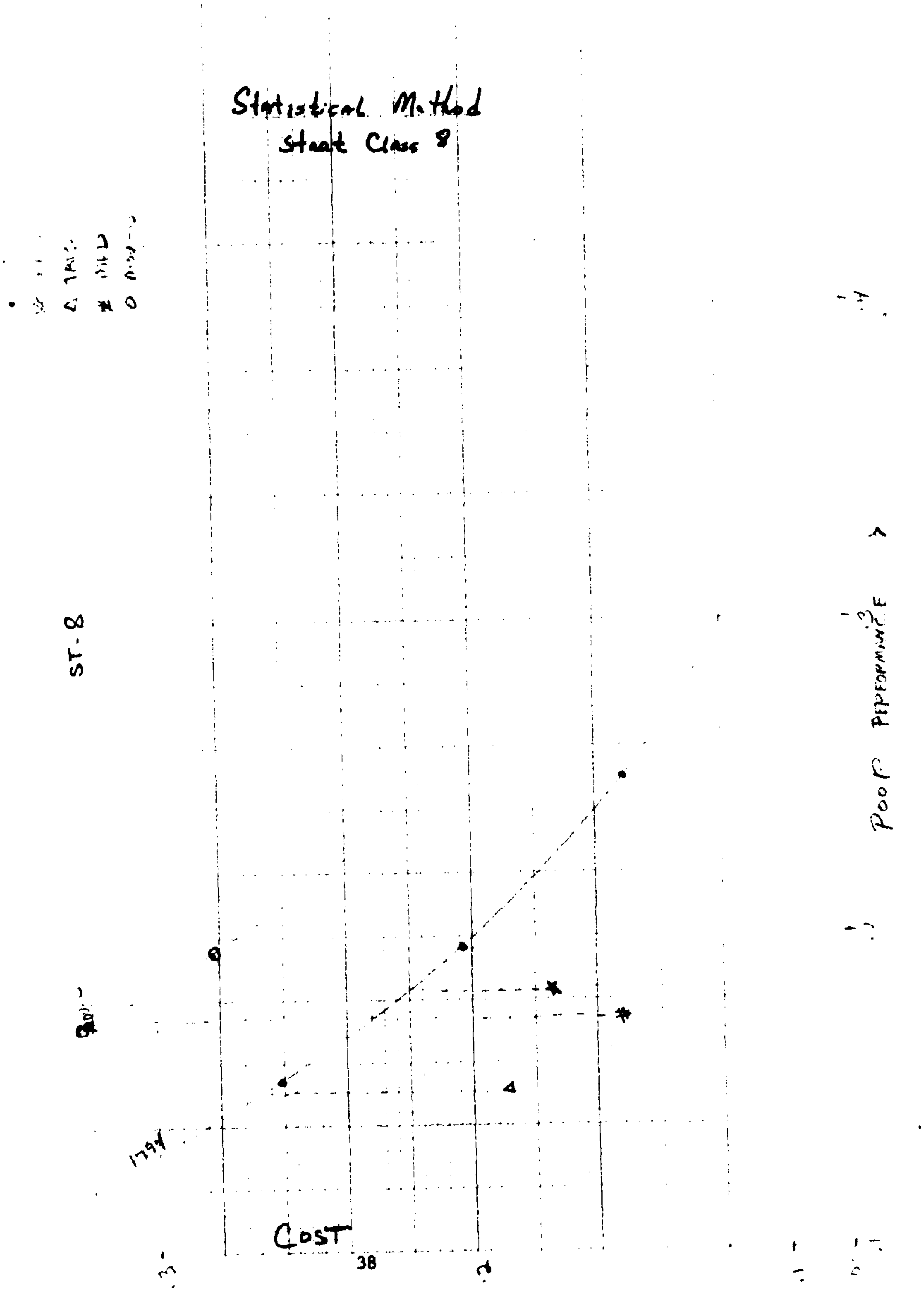
STRAT CLASSES: ST9, ST8, ST7

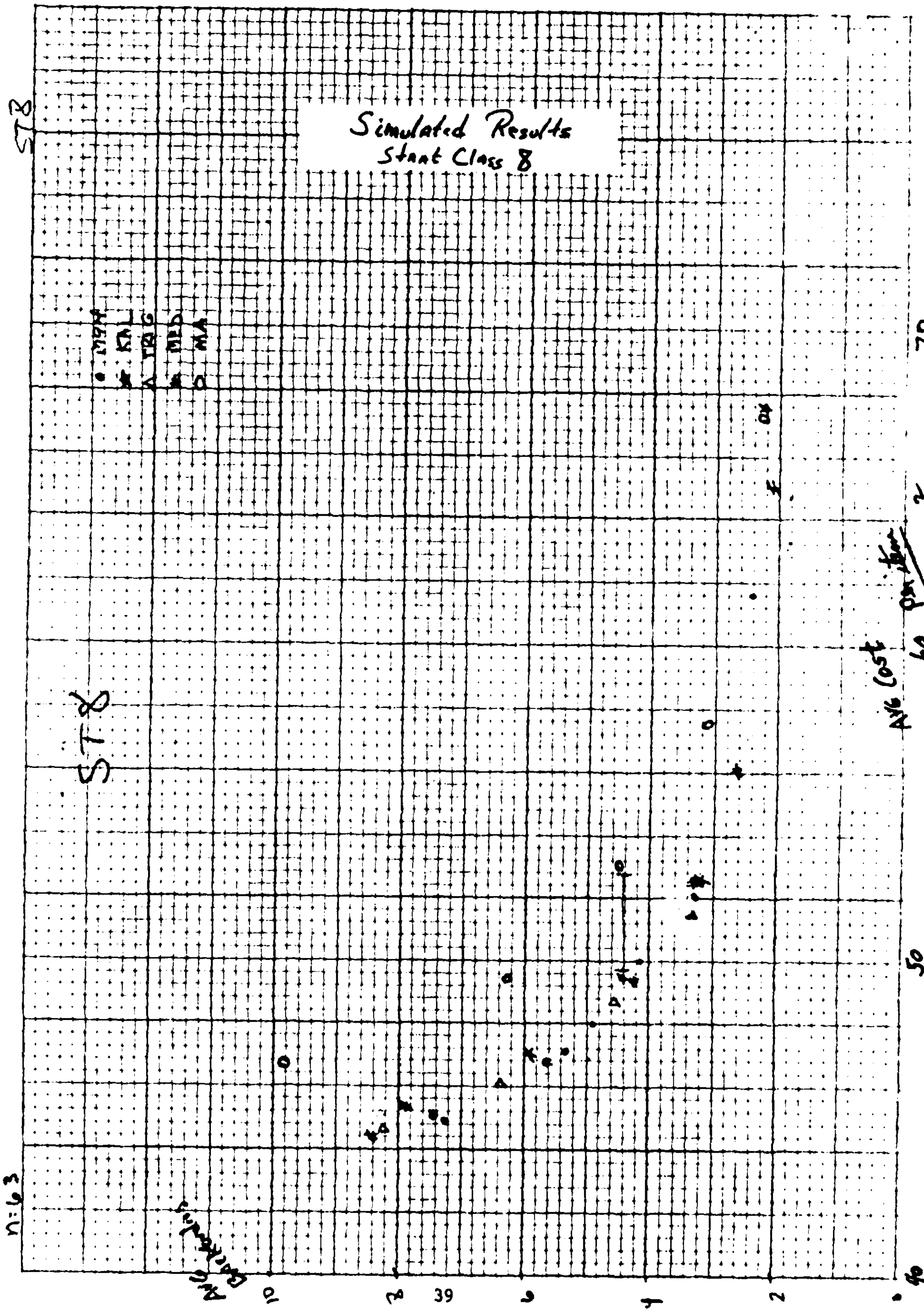
Statistical Method
Strat Class 9




Simulated Results
Strat Class 9
ST9
(All Items)
N=112
KIL
TRIG
AED
MA
Avg Cost
11000
11500
12000
12500
13000
13500
14000
14500
7.0
6.5
6.0
5.5
5.0
4.5
4

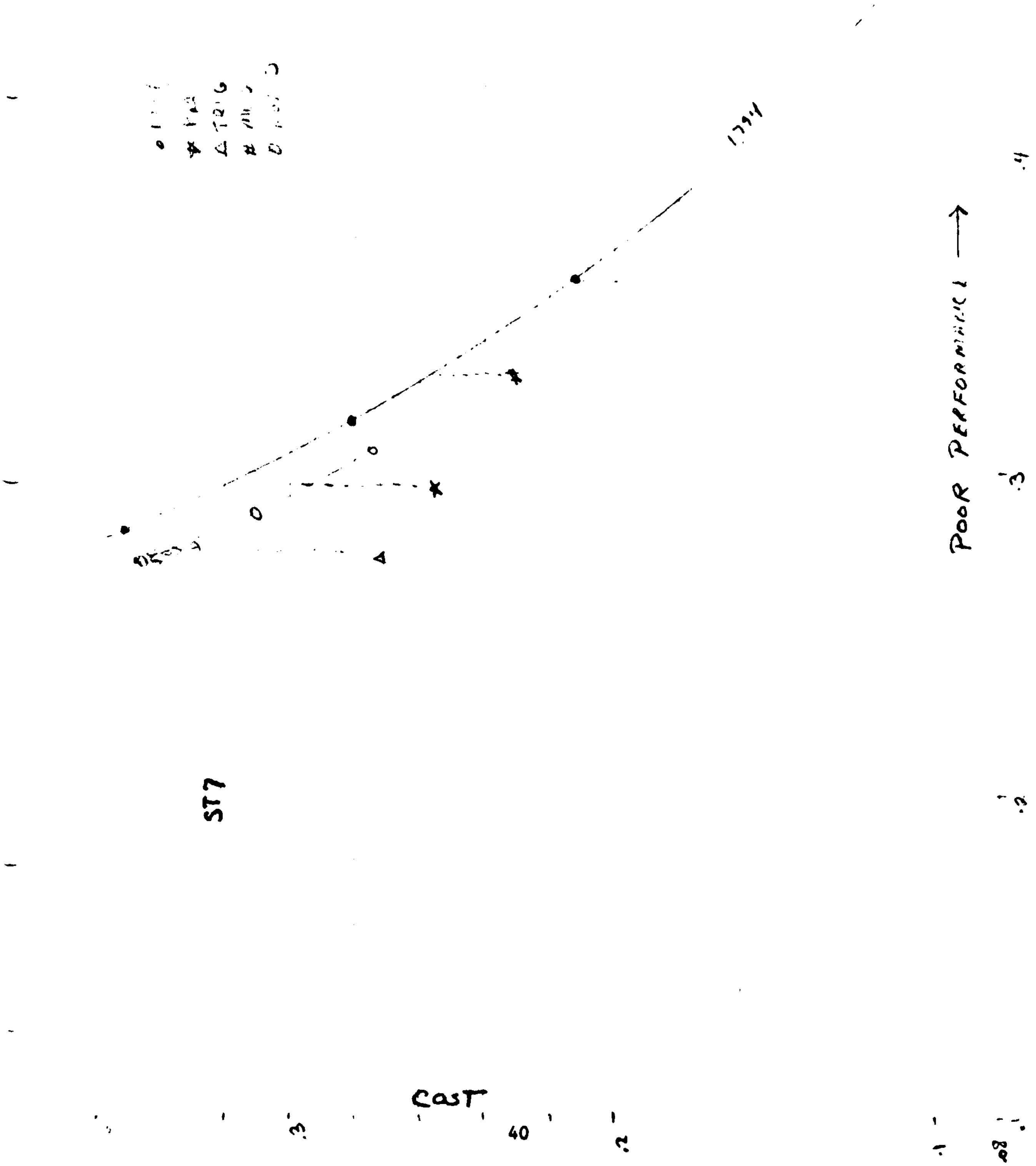
Statistical Method
Staat Class 8

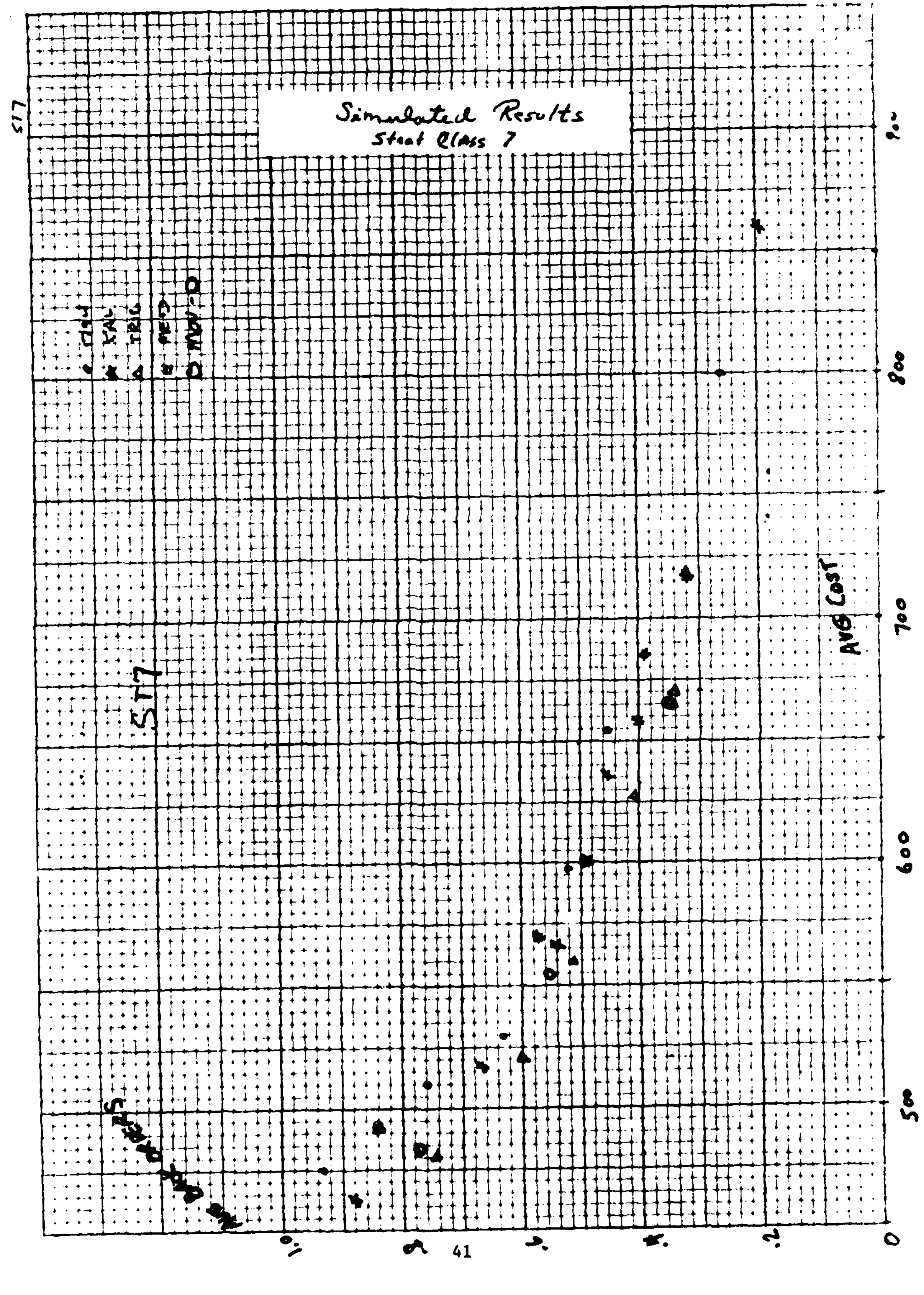




Simulated Results
Stant Class 8
ST8
n:63
70
60
50
40
10
3
2

Statistical Method
Strut Class 7




Simulated Results
Strat Class 7
ST7
KAL
TRIG
AVG COST
500
600
700
800
0

APPENDIX A11

STATISTICAL ERROR MEASURES USING A LOW DEMAND DATA BASE

Statistical Error Measures

Forecast Method	ERROR (BIAS)		RELATIVE ERROR		MAD		REL MAD		REL MSE	
	1st qtr	PLt	1st qtr	PLt	1st qtr	PLt	1st qtr	PLt	1st qtr	PLt
1799	.366	2.156	1.589	1.795	1.43	3.361	2.73	2.397	29.70	25.529
KAL	1.956	6.67	7.589	7.46	2.843	7.45	8.418	7.75	168.00	152.43
TRIG	.847	3.35	3.112	3.188	1.853	4.38	4.129	3.63	34.82	27.459
MOD	-.344	-.026	-.181	.078	.895	1.66	1.116	.907	6.57	3.22
MWD	.487	2.644	1.806	2.06	1.55	3.82	2.92	2.64	35.11	32.95
ZERO	-.697	-1.197	-.748	-.55	.69	1.197	.748	.54	5.80	1.197

Data Base = ST1

N = 335

LOW DEMAND ITEMS

REFERENCES

1. Cohen, Martin, "Alpha 4140.39 Simulator," AMC Inventory Research Office, May 1973, AD762348.

2. Cohen, Martin, "Demand Forecasting With Program Factors," AMC Inventory Research Office, September 1975, ADA017858.

3. Gotwals, E. P. & A. Hutchison, "Use of End Item Age in Demand Forecasting," DRC Inventory Research Office, December 1976, ADA037585.

4. Gotwals, E. P. and D. Orr, "Integrated Forecasting Techniques for Secondary Item Classes - Part 1, Active Items," US Army Inventory Research Office, September 1980, ADA097024.

5. Gotwals, E. P., "Integrated Forecasting Techniques for Secondary Item Classes - Part II, Inactive Items," US Army Inventory Research Office, September 1980, ADA097025.

6. Meng, Daniel, "Application of Dynamic Mean to Demand Forecast," System Analysis Division, Tank Automotive Materiel Readiness Command, September 1977, Report No. SA-77-11.

7. Orr, Donald, "Demand Forecasts Using Process Models and Item Class Parameters: Application of Ancillary Variables," DRC Inventory Research Office, April 1976, ADA026081.

8. Orr, Donald, "Kalman and Moving Average Filters for Forecasting: Systematics of Demand Processes and Extension," DRC Inventory Research Office, October 1976, ADA032476.

9. Orr, Donald and Alan Kaplan, "Data Enrichment for Simulation," Simulation, June 1976, 177-183.

<u>DISTRIBUTION</u>

COPIES	
1	Deputy Under Sec'y of the Army, ATTN: Office of Op Resch
	Headquarters, US Army Materiel Development & Readiness Command
1	DRCPS-P, ATTN: Ms. Lamb
1	DRCMM-R
1	DRCMM-S
1	DRCMM-RS
1	DRCMM-M
1	DRCMM-E
1	DRCRE
1	Dep Chf of Staff for Logistics, ATTN: DALO-SMS, Pentagon, Wash., DC 20310
1	Commandant, US Army Logistics Mgt Center, Ft. Lee, VA 23801
1	Office, Asst Sec'y of Defense, ATTN: MRA&L-SR, Pentagon, Wash., DC 20310
2	Defense Logistics Studies Info Exchange, DRXMC-D
10	Defense Technical Information Center, Cameron Station, Alexandria, VA 22314
	Commander, USA Armament Materiel Readiness Cmd, Rock Island, IL 61299
1	ATTN: DRSAR-MM
1	ATTN: DRSAR-SA
	Commander, USA Communications-Electronics Cmd, Ft. Monmouth, NJ 07703
1	ATTN: DRSEL-MM
1	ATTN: DRSEL-PL-SA
	Commander, USA Missile Command, Redstone Arsenal, AL 35898
1	ATTN: DRSMI-S
1	ATTN: DRSMI-D
	Commander, USA Troop Support & Aviation Materiel Readiness Command, St. Louis, MO 63120
1	ATTN: DRSTS-SPF
1	ATTN: DRSTS-SPS
1	ATTN: DRSTS-BA
	Commander, US Army Tank-Automotive Command, Warren, MI 48090
1	ATTN: DRSTA-F
1	ATTN: DRSTA-S
1	Commander, US Army Armament Research & Development Cmd, ATTN: DRDAR-SE, Dover, NJ 07801
1	Commander, US Army Aviation Research & Development Cmd, 4300 Goodfellow Blvd, St. Louis, MO 63120
1	Commander, US Army Aviation Research & Development Cmd, ATTN: DRDAV-BD (Marl Glenn), 4300 Goodfellow Blvd, St. Louis, MO 63120
1	Commander, US Army Electronics Research & Development Command, ATTN: DRDEL-ST-SA, Rm 2D315, Bldg 2700, Ft. Monmouth, NJ 07703
1	Commander, US Army Mobility Equipment Research & Development Cmd, ATTN: DRDME-O, Ft. Belvoir, VA 22060
1	Commander, US Army Natick Research & Development Command, ATTN: DRXNM-O, Natick, MA 01760
1	Commander, US Army Logistics Center, Ft. Lee, VA 23801
1	Commander, US Army Logistics Evaluation Agency, New Cumberland Army Depot, New Cumberland, PA 17070

COPIES	
1	Commander, US Army Depot Systems Command, Chambersburg, PA 17201
1	Commander, US Air Force Logistics Cmd, WPAFB, ATTN: AFLC/XRS, Dayton, Ohio 45433
1	US Navy Fleet Materiel Support Office, Naval Support Depot, Mechanicsburg, PA 17055
1	Mr. James Prichard, Navy SEA Systems Cmd, ATTN: PMS 3061, Dept of US Navy, Wash., DC 20362
1	George Washington University, Inst of Management Science & Engr., 707 22nd St., N.W., Wash., DC 20006
1	Naval Postgraduate School, ATTN: Dept of Opns Anal, Monterey, CA 93940
1	Air Force Institute of Technology, ATTN: SLGQ, Head Quantitative Studies Dept., Dayton, OH 43433
1	US Army Military Academy, West Point, NY 10996
1	Librarian, Logistics Mgt Inst., 4701 Sangamore Rd., Wash.,DC 20016
1	University of Florida, ATTN: Dept of Industrial Systems Engr., Gainesville, FL 32601
1	RAND Corp., ATTN: S. M. Drezner, 1700 Main St., Santa Monica, CA 90406
1	US Army Materiel Systems Analysis Activity, ATTN: DRXSY-CL, Aberdeen Proving Ground, MD 21005
1	Commander, US Army Logistics Center, ATTN: Concepts & Doctrine Directorate, Ft. Lee, VA 23801
1	ALOG Magazine, ATTN: Tom Johnson, USALMC, Ft. Lee, VA 23801
1	Commander, USDRC Automated Logistics Mgt Systems Activity, P.O. Box 1578, St. Louis, MO 63188
1	Director, DARCOM Logistics Systems Support Agency, Letterkenny Army Depot, Chambersburg, PA 17201
1	Commander, Materiel Readiness Support Activity, Lexington, KY 40507
1	Director, Army Management Engineering Training Agency, Rock Island Arsenal, Rock Island, IL 61299
1	Defense Logistics Agcy, ATTN: DLA-LO, Cameron Sta, Alexandria, VA 22314
1	Dep Chf of Staff (I&L), HQ USMC-LMP-2, ATTN: LTCJ Sonneborn, Jr., Wash., DC 20380
1	Commander, US Army Depot Systems Command, Letterkenny Army Depot, ATTN: DRSDS-LL, Chambersburg, PA 17201
1	Logistics Control Activity, Presidio of San Francisco, CA 94120
1	HQ, Dept of the Army, (DASG-HCL-P), Wash., DC 20314
1	Operations Research Center, 3115 Etcheverry Hall, University of California, Berkeley, CA 94720
1	Dr. Jack Muckstadt, Dept of Industrial Engineering & Operations Research, Upson Hall, Cornell University, Ithaca, NY 14890
1	Dept of Industrial Engineering, University of Michigan, Ann Arbor, MI 48104
1	Scientific Advisor, ATCL-SCA, Army Logistics Center, Ft. Lee, VA 23801
1	Prof Robert M. Stark, Dept of Stat & Computer Sciences, University of Delaware, Newark, DE 19711
1	Prof E. Gerald Hurst, Jr., Dept of Decision Science, The Wharton School, University of Penna., Phila., PA 19104
1	Logistics Studies Office, DRXMC-LSO, ALMC, Ft. Lee, VA 23801
1	Procurement Research Office, DRXMC-PRO, ALMC, Ft. Lee, VA 23801
1	Dept of Industrial Engr. & Engr. Management, Stanford University, Stanford, CA 94305

COPIES	
1	Commander, US Army Communications Command, ATTN: Dr. Forry, CC-LOG-LEO, Ft. Huachuca, AZ 85613
1	Prof Harvey M. Wagner, Dean, School of Business Adm, University of North Carolina, Chapel Hill, NC 27514
1	DARCOM Intern Training Center, Red River Army Depot, Texarkana, Tx 75501
1	Prof Leroy B. Schwarz, Dept of Management, Purdue University, Krannert Bldg, West Lafayette, Indiana 47907
1	US Army Training & Doctrine Command, Ft. Monroe, VA 23651
1	US General Accounting Office, ATTN: Mr. J. Morris, Rm 5840, 441 G. St., N.W., Wash., DC 20548
1	Operations & Inventory Analysis Office, NAVSUP (Code 04A) Dept of Navy, Wash., DC 20376
1	US Army Research Office, ATTN: Robert Launer, Math. Div., P.O. Box 12211, Research Triangle Park, NC 27709
1	Prof William P. Pierskalla, 3641 Locust Walk CE, Philadelphia, PA 19104
1	US Army Materiel Systems Analysis Activity, ATTN: DRXSY-MP, Aberdeen Proving Ground, MD 21005
	Air Force Logistics Management Center, Gunter Air Force Station, AL 36144
1	ATTN: AFLMC/LGY
1	ATTN: AFLMC/XRP, Bldg. 205
1	Engineer Studies Center, 6500 Brooks Lane, Wash., DC 20315
1	US Army Materiel Systems Analysis Activity, ATTN: Mr. Herbert Cohen, DRXSY-MP, Aberdeen Proving Ground, MD 21105
1	Dr. Carl Weisman, C.A.C.I., 1815 North Fort Myer Drive, Arlington,VA 22209
1	Commander, US Army Missile Cmd, ATTN: Ray Dotson, DRSMI-DS, Redstone Arsenal, AL 35898
1	Prof Peter Kubat, Graduate School of Mgmt, University of Rochester, Rochester, NY 14627
1	Prof Barney Bissinger, The Pennsylvania State University, Middletown, PA 17057
1	Prof Gary D. Scudder, Mgmt Sciences Dept, College of Bus Adm, University of Minnesota, 271 19th Ave., South, Minneapolis, MN 55455